

University of Groningen

Fair infinite lotteries

Wenmackers, Sylvia; Horsten, Leon

Published in:
Synthese

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2010

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Wenmackers, S., & Horsten, L. (2010). Fair infinite lotteries. *Synthese*.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Fair Infinite Lotteries *

Sylvia Wenmackers & Leon Horsten

September 1, 2010

Abstract

This article discusses the question how the concept of a fair finite lottery can best be extended to denumerably infinite lotteries. Techniques and ideas from non-standard analysis are brought to bear on the problem.

1 Introduction: from the finite to the transfinite

Although a fair lottery over the natural numbers may not objectively exist, we have strong intuitions about it. We may even give quantitative answers to questions such as: “What is the probability that the winning number is odd?” We may answer “Fifty percent”, by considering the probability of even and odd numbers in finite lotteries.

When forming an image of infinite mathematical objects, we rely on our experience with finite objects [Lavine 1995]. More often than not it is impossible to construct or discover an infinite counterpart of a finite concept that fulfils *all* our intuitions concerning the former. In such cases we have to choose which of those intuitions is most dear to us and weaken or give up at least one other. An example from classical mathematics would be the assignment of cardinalities to infinite sets by Cantor: he took the existence of a one-to-one correspondence between sets as the guiding principle for assigning equal sizes to them, but had to give up the intuition that the whole is always larger than the parts. If the infinite was like the finite in every relevant respect, then it would not be so interesting. Giving up some of our intuitions and tacit assumptions is just the price we have to pay if we want to study a new object. So it is with infinite lotteries. We will have to give up some of our intuitions governing finite lotteries. The question is: which ones, and to what extent.

*Thanks to Hannes Leitgeb, Vieri Benci, and Richard Pettigrew, for valuable discussions and comments. The contribution of SW was supported by the Odysseus Grant *Formal Epistemology: Foundations and Applications* funded by the Research Foundation Flanders (FWO-Vlaanderen). The contribution of LH was in part supported by the AHRC project *Foundations of Structuralism* (AH/H001670/1).

In this paper, we consider a fair lottery in which exactly one winner is randomly selected from a countably (or denumerably) infinite set of tickets. We intend to find a description of such a lottery that is mathematically sound and philosophically adequate, by examining our intuitions governing finite lotteries. It may be objected that since there are no infinite lotteries in reality, it is not clear how we can have any intuitions about the concept. Real world lotteries are always finite, but —no matter how the drawing is realized— the outcome can never be guaranteed to be random. Therefore even the idea of a truly fair n -ticket lottery, where n is some finite natural number, is a highly idealized concept, but one that is useful in analyzing a broad range of practical situations. Allowing the lottery to have an infinite number of possible outcomes is an additional idealization. The idea of a fair lottery on the natural numbers occurs in probabilistic number theory [Tenenbaum 1995] and may be a useful approximation for large lottery-like phenomena. This problem also goes by the name of ‘de Finetti’s lottery’ [Bartha 2004] and ‘God’s lottery’ [McCall & Armstrong 1989]. Although we will never be confronted with a lottery consisting of an infinite set of tickets in reality, it is valid to ask what probability can rationally be assigned to a ticket.

In the subjectivistic approach to probability, it has been argued (for instance by Ramsey [Ramsey 1926] and de Finetti [de Finetti 1974]) that our subjective probability assignments can only be rational if they agree with Kolmogorov’s laws of probability [Kolmogorov 1933]. Within Kolmogorov’s axiomatization however, there simply is no description available for a fair countably infinite lottery. To describe this case, we have to formulate new axioms or at least change one of the assumptions or axioms of Kolmogorov’s system. One solution, advocated by de Finetti [de Finetti 1974], is to relax the requirement of countable additivity to finite additivity. In this article, we will develop a different approach. We propose to replace the co-domain (or range) of the probability measure by a non-standard set: this allows us to assign a non-zero, infinitesimal probability to single tickets. In contrast to de Finetti’s solution, there will be a sense in which the probability of a countably infinite union of events supervenes on the probabilities of the individual events. This sense is captured by an additivity principle that is a close analogue of the usual assumption of countable additivity.

Structure of the paper. In section 2, we examine our intuitions concerning finite and infinite lotteries. In section 3 we review asymptotic density and a generalization thereof by means of which a finitely additive, real-valued probability measure can be obtained on the full power set of the natural numbers. In section 4, we introduce some central concepts of non-standard analysis. In section 5, we construct a hyperrational-valued probability measure (based on the concept of numerosity) and show that it

is hypercountably additive rather than countably additive. In section 6, we compare our hyperrational approach with the real-valued solution based on (generalized) asymptotic density, and with a hyperfinite lottery. In the 7th and last section, we review the most salient features of the proposed hyper-rational description of a fair, countably infinite lottery.

Notation. Throughout this paper we take \mathbb{N} to be the set of strictly positive integers, $\{1, 2, 3, \dots\}$. *Even* will denote the set of even natural numbers and *Odd* that of the odd natural numbers. We use $\langle \rangle$ -brackets to indicate an ω -sequence; if only one element is given between the brackets, it will be the general element at a position $n \in \mathbb{N}$. We abbreviate ‘non-standard analysis’ as NSA, ‘finite additivity’ and ‘finitely additive’ with FA, and ‘countable additivity’ and ‘countably additive’ with CA. The ‘H’ in HFA and HCA adds the prefix ‘hyper-’ to the former abbreviations.

2 Intuitions concerning lotteries

First we need to be precise about what we mean exactly with a finite lottery, give a mathematical description of it, and make our intuitions about it explicit. Subsequently we investigate to what extent these intuitions carry over to the infinite case, and what has to be changed in the mathematical description to maximize the intuitive appeal of it.

2.1 Finite lotteries

Probability measure By a finite lottery we mean a process that assigns exactly one winner among a discrete set of tickets in a fair way. By fair we mean that each ticket initially has the same probability of winning. So this process can be modeled by a uniform, discrete function (given below in eq.(1)), which fulfills all of Kolmogorov’s axioms for probability measures [Kolmogorov 1933].

The sample space is the set of tickets. The tickets may be numbered, but they need not be: they may be characterized by other symbols with no apparent order. Since the tickets are finite in number, say $n \in \mathbb{N}$, they can be labeled with an initial segment of the natural numbers, and this set of numbers $\{1, \dots, n\}$ may be used as the sample space instead.

The event space is a σ -algebra¹ which contains all combinations of tickets to which we can assign a probability. Since we may do so for any possible combination of tickets, the event space is the powerset $\mathcal{P}(\{1, \dots, n\})$ of the sample space. The probability values for an n-ticket lottery form the set

¹An algebra is a family of subsets of the sample space that contains the sample space itself and is closed under complementation and finite unions; a σ -algebra is closed under denumerable unions on top of that.

$\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\}$. The co-domain of a probability measure on a finite lottery with an unspecified number of tickets is therefore the set of all rational numbers in the $[0, 1]$ interval: $[0, 1]_{\mathbb{Q}} = [0, 1] \cap \mathbb{Q}$. The probability measure is given by the function:

$$\begin{array}{ccc} P_n : \mathcal{P}(\{1, \dots, n\}) & \rightarrow & [0, 1]_{\mathbb{Q}} \\ A & \mapsto & \frac{\#(A)}{n} \end{array} \quad (1)$$

where $\#$ is the counting function, that maps a finite set to its number of elements (finite cardinality). Thus, P_n is a counting measure normalized by the total number of elements in the sample space. $\#$ is the prototype of a CA measure, a property that we will employ in the proof of the additivity of the non-standard measure that we will construct.

Thus, P_n is CA too, but only in a trivial sense: for each countable family of disjoint subsets $\{A_1, A_2, A_3, \dots\}$ of the sample space, there will be a finite value $k \in \mathbb{N}$ such that for all $m \geq k$, A_m is an empty set. Hence, in the countable sum $\sum_{m \in \mathbb{N}} P_n(A_m)$, all terms with $m \geq k$ will be zero and CA reduces to FA in this case.²

Intuitions Now we list our intuitions governing a finite lottery. Some of these may seem highly related, and they are—at least in the finite case—, but they need not be in the infinite case, so we name them separately:

FAIR The lottery is fair.

ALL Every ticket has a probability of winning.

SUM The probability of a combination of tickets can be found by summing the individual probabilities.

LABEL The labelling of the tickets is neutral with respect to the outcome.

The assumption FAIR embodies the thought that one ticket does not have a higher probability than any other one: a fair lottery is governed by equiprobability. This can only be implemented by the formal requirement that the associated probability function is uniform.

The assumption ALL can only be implemented by the requirement that the probability of any possible combination of tickets is defined. In other words, the probability function must be defined over the whole power set of the event space.

The assumption LABEL is motivated by the intuition that labelling is no more than a convention inspired by the need for referring to specific tickets.

² k cannot be fixed in general, but depends on the family. Although at most n sets can be non-empty, k may need to be larger than n , since a family of sets can consist of many empty sets ‘in between’ the non-empty ones.

It is implemented by requiring of the associated probability function that it is invariant under permutations of the domain.

The assumption SUM is motivated by the intuition that the probability of a set containing the winning number supervenes on the chances of winning that accrue to the individual tickets. The usual assumption of countable additivity (CA, sometimes also called σ -additivity) is one attempt of making the intuition that is encapsulated by SUM precise. We will argue, however, that this is not the right way to do it in this case. In other words, we will argue that the implementation of SUM is not as straightforward an affair as is commonly thought.

The constraints FAIR, ALL, and SUM jointly entail that every point event must be assigned a non-zero probability. Thus in the context of infinite lotteries the so-called principle of Regularity holds. This is not to say, however, that this principle must hold in all probabilistic scenarios.³

These assumptions motivate the standard account of finite lotteries, and they are jointly satisfied by the standard description given by P_n . Let us now briefly survey how these assumptions fare in the context of infinite lotteries.

2.2 Infinite lotteries

The infinite counterpart of a finite lottery that we are interested in here is an infinite, denumerable lottery, in particular a lottery that has \mathbb{N} as its sample space.

For a start, suppose that we are very keen on the intuitions ALL and SUM, and that SUM is formally cashed out as CA. There are (uncountably) many probability distributions that satisfy these two constraints. However, it is easy to see that all of them violate FAIR. But the assumption FAIR simply is non-negotiable. The intuition of fairness is absolutely central to our concept of a lottery. Whereas real world lotteries may never be completely fair, we are considering ideal lotteries. Indeed, when one considers infinite lotteries at all, one is leaving the real world behind anyway.

The assumption LABEL might seem reasonable at first blush, but as a consequence of Cantor's theory of infinite cardinalities, it will have to be abandoned. Every infinite subset of \mathbb{N} is in one-to-one onto correspondence with every other infinite subset of \mathbb{N} . So if we insist on invariance under permutation, then every infinite subset of \mathbb{N} will receive the same probability. This immediately leads to a contradiction with the laws of probability.⁴ Thus, whereas for finite sample spaces the labelling of the point events is immaterial, in the infinite case it is of the essence. Giving up LABEL

³Skyrms and Lewis defend Regularity in general [Skyrms 1980, Lewis 1980], whereas Williamson, Hájek and Easwaran argue that it cannot hold in all cases (even when infinitesimal probability values are allowed) [Williamson 2007, Hájek, Easwaran].

⁴By considering *Even* and *Odd* on the one hand, and their union \mathbb{N} on the other hand, we see that all three sets have the same measure, so FA fails.

admittedly gives rise to a feeling of discomfit. But if we have a naked choice between giving up some of the laws of probability on the one hand, and giving up LABEL on the other hand, then we should surely take the second option.⁵

The assumption ALL seems negotiable to some extent. For one thing, it is a well-known consequence of the Axiom of Choice that there is no probability measure on the whole of $\mathcal{P}(\mathbb{R})$ [Truss 1997, chapter 11]. If there are no probability measures on $\mathcal{P}(\mathbb{R})$, then it should perhaps come as no surprise that there are no *satisfactory* probability measures on $\mathcal{P}(\mathbb{N})$ in the context of infinite lotteries. In any event, we will require of any solution to the infinite lottery problem that to the extent that we have strong intuitions about probabilities of a subset $A \subseteq \mathbb{N}$, the solution takes the probability of A to be defined and in agreement with our intuitions —or else the solution will have to contain a winning story about why our intuitions are mistaken.

Like ALL, the assumption SUM is not so easy to assess. It is well-known that there is no uniform probability function on $\mathcal{P}(\mathbb{N})$ that is CA [de Finetti 1974]. So if we want to preserve FAIR and ALL, and insist on SUM, then we must make it precise in a way that is different from CA.

3 Asymptotic density: real-valued probability with finite additivity

3.1 Limiting relative frequency

Although all infinite subsets of the natural numbers have the same cardinality, there are ways to discriminate the ‘size’ of *Even* from that of \mathbb{N} for instance. In number theory, the asymptotic density *ad* (or natural density) of a subset A of \mathbb{N} is defined as follows (e.g. in [Tenenbaum 1995, p. 270]):

$$ad(A) = \lim_{n \rightarrow \infty} \frac{\#(A \cap \{1, \dots, n\})}{n} \quad \text{if the limit exists} \quad (2)$$

Asymptotic density captures the idea that a lottery over \mathbb{N} is obtained from a finite lottery (eq. (1)) in the limit of the number of tickets, n , going to infinity. Thus, for a set A that has a defined asymptotic density $ad(A) = \lim_{n \rightarrow \infty} P_n(A)$. Since \mathbb{Q} is not closed under the limit-operation, the co-domain of *ad* is the $[0, 1]$ -interval of the real numbers, rather than $[0, 1]_{\mathbb{Q}}$ as in the finite case.

Asymptotic density gives rise to a FAIR probability assignment. However, it fails ALL: since asymptotic density is not defined for all subsets of \mathbb{N} , it cannot be introduced as a measure with $\mathcal{P}(\mathbb{N})$ as its domain [Tenenbaum 1995].

⁵For a discussion of the role of LABEL in assigning probabilities in the context of the realism debate, see [Douven et al 2010].

The collection of all subsets of \mathbb{N} that do possess asymptotic density is not closed under intersection and union, so it does not form an algebra.

3.2 A generalization

It is possible to extend asymptotic density to a measure that assigns a value to all subsets of \mathbb{N} ([Schurz & Leitgeb 2008]). This requires a generalization of the limit-concept, that assigns a value to all bounded —convergent and non-convergent— sequences: the Hahn-Banach limit (HB – lim) is a real-valued generalization, that equals the value of the classical limit for convergent sequences.⁶ For any $A \in \mathcal{P}(\mathbb{N})$ the sequence $\langle \frac{\#(A \cap \{1, \dots, n\})}{n} \rangle$ is bounded (by 1). Therefore, the Hahn-Banach limit of this sequence is defined on all of $\mathcal{P}(\mathbb{N})$, giving rise to the probability measure P_{ad} :

$$\begin{aligned} P_{ad}: \mathcal{P}(\{1, \dots, n\}) &\rightarrow [0, 1]_{\mathbb{R}} \\ A &\mapsto \text{HB} - \lim \frac{\#(A \cap \{1, \dots, n\})}{n} \end{aligned} \quad (3)$$

The great selling point of this construction is that probabilities are aligned with limiting relative frequencies whenever these are defined. But again, the resulting measure is only FA [Schurz & Leitgeb 2008]. So even though ALL can be obtained, SUM does not hold.

P_{ad} can be thought of as giving precise content to a suggestion by de Finetti: by weakening the requirement for CA to FA, all other axioms of classical probability theory can be saved. This solution to the problem of the fair infinite lottery was advocated by de Finetti [de Finetti 1974]. It allows us to assign probability values to infinite subsets of \mathbb{N} that correspond well to our intuitions (such as a probability 1/2 for *Even* as well as for *Odd*). The solution does come with a major drawback, however. It amounts to giving up on the intuition that the chance of a ticket from an infinite set winning is an infinite sum over the chances of individual tickets from the set winning (SUM). As a result, regularity fails, so we also have to give up on the intuition that each ticket has a non-zero chance of winning.

4 Infinitesimals

With FA, we can save FAIR, ALL and a finite version of SUM. But it seems odd that the measure of singletons can be exactly 0 while the measure of their union is 1. Now the question is: can we do better? So far, we have only looked at real-valued probability functions, and we have seen that in that framework the answer is ‘no’. Now we will reconsider the question. It seems that we can do better indeed, by assigning an infinitesimal probability to the singletons, rather than 0. In the finite case, the probability of a singleton is 1 divided by the number of elements in the sample space. Since our

⁶We will give a definition of HB – lim in the non-standard framework in paragraph 4.5.

sample space is \mathbb{N} , with an infinite number of elements, we should assign the inverse of an infinite number to the singletons. Cantor's cardinalities are not suitable for this, since they have no inverse. Non-standard analysis provides a consistent way of working with unbounded or infinite numbers, which do have an inverse: infinitesimals. So, non-standard analysis allows us to have a different co-domain for the probability measure: ${}^*[0, 1]{}^*\mathbb{Q}$, which is the unit interval within ${}^*\mathbb{Q}$, a non-standard extension of \mathbb{Q} .

We will give a short overview of some essential concepts in non-standard analysis (NSA) [Robinson 1966]. We do not aim at completeness here, but restrict our attention to those ingredients that we will need in the course of this paper. The information of this section is mainly based on [Cutland 1983] and [Benci et al 2006a].

4.1 The star-map and Transfer

All approaches to NSA need a tool that maps any standard object, A (which can be a number, set, function, \dots), to its unique non-standard counterpart or hyperextension, *A . This function, called the star-map, should preserve a large class of properties, which is ensured by the Transfer principle. In axiomatic approaches to NSA, the Transfer principle is stipulated as an axiom; in other approaches such as the ultrafilter-construction that we will employ here, it is a theorem.

An important example is the star-map of an ω -sequence. This is a hypersequence, i.e., a sequence that takes values on all of ${}^*\mathbb{N}$, which is a non-standard extension of \mathbb{N} . We introduce the notation $\langle\langle \cdot \rangle\rangle$ to distinguish a hypersequence from a standard sequence; if only one element is given between the brackets, it will be the general element at position N , with $N \in {}^*\mathbb{N}$. (We will encounter the star-map of a sequence of sets in eq. (23).)

The star-map can be obtained using free ultrafilters; this requires the introduction of equivalence classes determined by a free ultrafilter.

4.2 Equivalence classes based on a free ultrafilter

Consider the set of ω -sequences on a general set X (or the set $X^{\mathbb{N}}$ of functions $\mathbb{N} \rightarrow X$). The idea is to interpret a whole sequence as one entity, be it a non-standard one. Even if X is a set of numbers, when $X^{\mathbb{N}}$ fails to form a field, it does not provide a useful number system. To this end, we need to make a choice of 'which positions in the sequence matter'. For instance, a difference in only finitely many positions should not matter. Fixing a free ultrafilter on the label set, \mathbb{N} , is a way to settle all these choices at once.⁷ A nice introduction to ultrafilters can be found in [Komjáth & Totik 2008].

⁷In general, non-standard analysis may be developed from considering a free ultrafilter on any infinite set, but in this paper we will always use \mathbb{N} as the index set.

A free (or non-principal) ultrafilter, \mathcal{U} , on \mathbb{N} is a collection of subsets of \mathbb{N} ($\mathcal{U} \subset \mathcal{P}(\mathbb{N})$), which fulfills four requirements:

1. $\emptyset \notin \mathcal{U}$
2. $(\forall A, B \in \mathcal{U}) \quad A \cap B \in \mathcal{U}$
3. $(\forall A \subset \mathbb{N}) \quad A \notin \mathcal{U} \Rightarrow \mathbb{N} \setminus A \in \mathcal{U}$ (ultra)
4. $(\forall A \subset \mathbb{N}) \quad A \text{ is finite} \Rightarrow \mathbb{N} \setminus A \in \mathcal{U}$ (free)

Two sequences are equivalent (or equal ‘almost everywhere’) with respect to a free ultrafilter just if the set of labels where their terms are exactly equal is an element of the filter:

$$(\forall \langle x_n \rangle, \langle y_n \rangle \in X^{\mathbb{N}}) \quad \langle x_n \rangle \approx_{\mathcal{U}} \langle y_n \rangle \Leftrightarrow \{n \mid x_n = y_n\} \in \mathcal{U} \quad (4)$$

We may define the equivalence class of a sequence $\langle x_n \rangle$ modulo the just defined equivalence relation, $[\langle x_n \rangle]_{\mathcal{U}}$, as follows:

$$(\forall \langle x_n \rangle \in X^{\mathbb{N}}) \quad [\langle x_n \rangle]_{\mathcal{U}} = \{\langle y_n \rangle \in X^{\mathbb{N}} \mid \langle y_n \rangle \approx_{\mathcal{U}} \langle x_n \rangle\} \quad (5)$$

The set of equivalence classes of sequences does provide a good basis for a number system and may be interpreted as *X , the hyperextension of X .

4.3 ${}^*\mathbb{N}$ and ${}^*\mathbb{Q}$

To illustrate how the equivalence class of sequences can provide the star-map of a set X , we consider the hyperextensions of $X = \mathbb{N}$ and $X = \mathbb{Q}$.

${}^*\mathbb{N}$ is defined as the set of equivalence classes (determined by the choice of a free ultrafilter on \mathbb{N}) of elements of $\mathbb{N}^{\mathbb{N}}$, $[\langle m_n \rangle]_{\mathcal{U}}$. The elements of ${}^*\mathbb{N}$ are called hypernatural numbers. The equivalence class of a constant sequence $[\langle m_n \rangle]$, with $m_n = m \in \mathbb{N}$ for all $n \in \mathbb{N}$, can be written in short as *m and may be identified with the standard natural number m : this embeds \mathbb{N} in ${}^*\mathbb{N}$. Because of the construction, ${}^*\mathbb{N}$ is called a sequential extension of \mathbb{N} : it consists of (equivalence classes of) ω -sequences of standard natural numbers.

For probability values, the sequential extension of \mathbb{Q} seems more appropriate than ${}^*\mathbb{N}$. The set of hyperreal numbers ${}^*\mathbb{Q}$ can be obtained in several ways. Since the set of integers, \mathbb{Z} , can be introduced as the closure of \mathbb{N} under subtraction, and the set of rational numbers, \mathbb{Q} , can be introduced as the fraction field of \mathbb{Z} , ${}^*\mathbb{N}$ can be extended similarly to ${}^*\mathbb{Z}$ and ${}^*\mathbb{Q}$ subsequently. Alternatively, by considering sequences of integer or rational numbers, ${}^*\mathbb{Z}$ and ${}^*\mathbb{Q}$ can be constructed using ultrafilters in a similar fashion as ${}^*\mathbb{N}$. (As was already mentioned, the star-map returns the hyperextension of any standard object, so also ${}^*\mathbb{R}$ and ${}^*\mathbb{C}$ can be obtained.)

4.4 \mathbb{R} as an approximation to ${}^*\mathbb{Q}$

To relate results of NSA to standard analysis, the standard part (or shadow) function, st , is a useful concept: st maps any non-standard number to the closest real value, which is uniquely determined. Clearly, ‘taking the standard part’ comes down to rounding up to infinitesimals.

For fair lotteries on finite sample spaces, the probabilities are fractions (rational numbers). For infinite sample spaces, probabilities can be associated with ω -sequences of rational values. Observe that both the construction of \mathbb{R} from Cauchy sequences and the above construction of ${}^*\mathbb{Q}$ start from ω -sequences of rational numbers. Only the rule by which the whole sequence is associated with a new —real or hyperrational— number differs. Thus, for the range of a fair infinite lottery we seem to have a choice between \mathbb{R} and ${}^*\mathbb{Q}$ (rather than between \mathbb{R} and ${}^*\mathbb{R}$).

If we use the standard part function on ${}^*\mathbb{Q}$, we can obtain any value of \mathbb{R} . Thus, from the non-standard viewpoint, real-valued probabilities can be seen as an approximation of hyperrational numbers. If the real values give satisfactory answers, they are all we need. If not, we may need to look at a more precise description in terms of hyperrationals.

4.5 Limits

Alpha-limit We may call ${}^*a_\alpha$ the ‘ideal value’ or the ‘alpha-limit’ of the sequence $\langle a_n \rangle$. It is equal to the value of the hypersequence $\langle\langle {}^*a_N \rangle\rangle$ at position $N = \alpha$ and also to the ultrafilter equivalence class of the sequence $\langle a_n \rangle$ [Benci & Di Nasso 2003a, p. 367]:

$${}^*a_\alpha = [\langle a_n \rangle]_{\mathcal{U}} \quad (6)$$

Classical limit The definition of the classical limit of a real-valued sequence, $\lim_{n \rightarrow \infty} a_n$, can be reformulated in NSA as follows [V  th 2007, p. 88]: $\langle a_n \rangle$ converges with limit $L \in \mathbb{R}$ if and only if

$$(\forall N \in {}^*\mathbb{N} \setminus \mathbb{N}) \quad {}^*a_N - L \text{ is infinitesimal} \quad (7)$$

In particular for $N = \alpha$, if $\lim_{n \rightarrow \infty} a_n = L$, then $st({}^*a_\alpha) = L$. By equation (6), we find that:

$$\lim_{n \rightarrow \infty} a_n = st\left([\langle a_n \rangle]_{\mathcal{U}}\right) \quad \text{if the limit exists} \quad (8)$$

Hahn-Banach limit Hahn-Banach limits (HB – lim) are a real-valued generalization of the limit that is defined for all bounded sequences and is equal to the value of the classical limit if the sequence converges. In NSA, the Hahn-Banach limit of a real-valued sequence $\langle a_n \rangle$ is defined as follows

[V  th 2007, p. 133]: $\langle a_n \rangle$ is bounded with $\text{HB} - \lim(a_n) = L \in \mathbb{R}$ if and only if

$$L = st \left(\sum_{N=H_0}^{H_1} h_N {}^*a_N \right) \quad (9)$$

for some $H_0, H_1 \in {}^*\mathbb{N} \setminus \mathbb{N}$ with $H_0 < H_1$ and for some internal sequence of hyperreals h_{H_0}, \dots, h_{H_1} such that $\sum_{N=H_0}^{H_1} h_N = 1$.

4.6 Internal and external objects

Not all non-standard objects are the image of some standard object by the star-map; those that are, are called ‘internal’, the others ‘external’. This distinction is important but not easy to understand immediately. As a first example, \mathbb{N} is external; in fact, every infinite countable set is external. It is also important to note that whereas the star-map preserves most set-operations, it does not do so for the powerset: the operation ${}^*\mathcal{P}$ returns only the internal subsets of a given set, and so, for any infinite standard set A : ${}^*(\mathcal{P}(A)) \subsetneq \mathcal{P}({}^*A)$. The probability function we will construct will also turn out to be external.

5 Hyperrational valued probability

Let us now construct a non-standard valued probability measure to describe a lottery on \mathbb{N} . If we want to find a function that is formally similar to the probability measure of a finite lottery, then we must find a way to ‘count’ finite as well as infinite subsets of \mathbb{N} and divide by the size it assigns to the whole sample space to normalize this function.

The construction proceeds in four steps. 1) Every subset of \mathbb{N} can be represented as an infinitely long bit string by considering its characteristic function. 2) Then we consider the sequence of partial sums of these bits. 3) By introducing a free ultrafilter on \mathbb{N} , we can interpret the whole partial sum sequence as one nonstandard (or hypernatural) number: its numerosity. 4) By a suitable normalization, we finally obtain a hyperrational-valued probability measure on $\mathcal{P}(\mathbb{N})$.

After the construction we show that numerosities and probabilities based on them are hypercountably additive (HCA). In fact, we shall see that it is perhaps more appropriate to call them hyperfinitely additive (HFA).

5.1 The construction

Step 1: Characteristic bit string First, consider the indicator function or characteristic function of a subset A of \mathbb{N} : it tests whether a natural number is in the set A or not, where a positive answer corresponds to 1 and a negative to 0.

$$\begin{aligned}\chi_A: \mathbb{N} &\rightarrow \{0, 1\} \\ n &\mapsto \begin{cases} 0 & \text{if } n \in \mathbb{N} \setminus A \\ 1 & \text{if } n \in A \end{cases}\end{aligned}\quad (10)$$

Now we can construct the function that assigns a characteristic bit string (CBS) to any subset of the natural numbers:

$$\begin{aligned}CBS: \mathcal{P}(\mathbb{N}) &\rightarrow \{0, 1\}^{\mathbb{N}} \\ A &\mapsto \langle \chi_A(1), \chi_A(2), \dots, \chi_A(n), \dots \rangle\end{aligned}\quad (11)$$

In shorthand notation, we can refer to a sequence by its n -th element only, so $CBS(A) = \langle \chi_A(n) \rangle$.

Step 2: Partial sums of characteristic bit strings Now we consider the sequence of partial sums of characteristic bit strings of a subset A of natural numbers, $SumCBS(A)$.

$$\begin{aligned}SumCBS: \mathcal{P}(\mathbb{N}) &\rightarrow \mathbb{N}^{\mathbb{N}} \\ A &\mapsto \langle S_n \rangle\end{aligned}\quad (12)$$

with

$$\begin{aligned}S_n &= \sum_{m=1}^n \chi_A(m) \\ &= \chi_A(1) + \chi_A(2) + \dots + \chi_A(n)\end{aligned}\quad (13)$$

So S_n has a value in $\{0, \dots, n\}$ for all n . Alternatively, S_n can be written in terms of the counting function $\#$:

$$S_n = \sum_{m=1}^n \chi_A(m) \quad (14)$$

$$= \sum_{m \in \mathbb{N}} \chi_{A \cap \{1, \dots, n\}}(m) \quad (15)$$

$$= \#(A \cap \{1, \dots, n\}) \quad (16)$$

The sequence $\langle S_n \rangle$ seems to ‘point to a value at infinity’. If we interpret this sentence in the framework of standard analysis, we should take the limit $n \rightarrow \infty$. For infinite sets A this results in $\lim_{n \rightarrow \infty} S_n = \infty$, which cannot be normalized. (If we consider the sequence $\langle \frac{S_n}{n} \rangle$ instead and its limit $n \rightarrow \infty$, we find asymptotic density again.) As we have seen however, in NSA we may alternatively interpret the whole sequence as one non-standard number.

Step 3: Numerosity as the equivalence class of the partial sum sequence The equivalence class under a free ultrafilter of a sequence of partial sums of characteristic bit strings is a hypernatural number that can be interpreted as the size of the corresponding set, called its *numerosity*.

$$\begin{aligned} num: \mathcal{P}(\mathbb{N}) &\rightarrow {}^*\mathbb{N} \\ A &\mapsto [\langle S_n \rangle]_{\mathcal{U}} \end{aligned} \quad (17)$$

with S_n as before.

Any finite set has a finite numerosity (value in \mathbb{N} , equal to $\#(A)$), whereas any infinite set has an infinite numerosity (value in ${}^*\mathbb{N} \setminus \mathbb{N}$). As an example, consider $A = \mathbb{N}$. In that case, $CBS(A) = \langle 1, 1, 1, \dots, 1, \dots \rangle$ and $num(A) = [\langle 1, 2, 3, \dots, n, \dots \rangle]_{\mathcal{U}}$, which is larger than any finite number. Thus we have shown that $num(\mathbb{N})$ is an element of ${}^*\mathbb{N} \setminus \mathbb{N}$. We may call this new number, $num(\mathbb{N})$, alpha (α).

Using equation (14), we can transform the definition for numerosity given by equation (17) in the following way:

$$\begin{aligned} num(A) &= [\langle S_n \rangle]_{\mathcal{U}} \\ &= [\langle \#(A \cap \{1, \dots, n\}) \rangle]_{\mathcal{U}} \\ &= [{}^*\#({}^*A \cap [\langle \{1, \dots, n\} \rangle]_{\mathcal{U}})]_{\mathcal{U}} \end{aligned} \quad (18)$$

for all subsets A of \mathbb{N} .⁸ Note that $[\langle \{1, \dots, n\} \rangle]_{\mathcal{U}}$ is a hyperfinite set, which means that there is an infinite hypernatural number, N , such that this set is equal to $\{1, \dots, N\}$. We can even be more precise, the hypernatural number N is equal to $[\langle 1, 2, 3, \dots, n, \dots \rangle]_{\mathcal{U}}$, which we will call α (following the terminology of Benci & Nasso). So we obtain:

$$num(A) = {}^*\#({}^*A \cap \{1, \dots, \alpha\}) \quad (19)$$

Remark how similar this form is to the numerator of the asymptotic density function in equation (2). Thus, equation (19) is very suggestive for a probability function: all we need to do is normalize it.

Step 4: Non-standard probability The construction is completed by normalization of the numerosity function. By dividing it by the numerosity of the sample space, $num(\mathbb{N}) = \alpha$, we can introduce the probability function of a lottery over \mathbb{N} in a form that is similar to the probability measure of a finite lottery (eq. (1)):

$$\begin{aligned} P_{num}: \mathcal{P}(\mathbb{N}) &\rightarrow {}^*[0, 1]_{{}^*\mathbb{Q}} \\ A &\mapsto \frac{num(A)}{\alpha} \end{aligned} \quad (20)$$

⁸The form on the second line makes it clear that our ultrafilter-based definition of the numerosity function is equal to that of the axiomatic approach developed in [Benci & Nasso 2003b]. In the existence proof given in that paper, numerosity is related to the ultrafilter-construction of NSA as follows: $num(A) = [\langle \#(A \cap \{1, \dots, n\}) \rangle]_{\mathcal{U}}$ [Benci & Nasso 2003b, p. 62], where \mathcal{U} is a free ultrafilter such that $\mathcal{U} = \{A \subseteq \mathbb{N} \mid \alpha \in {}^*A\}$ [Benci & Di Nasso 2003a, p. 374].

P_{num} takes values on the unit interval of ${}^*\mathbb{Q}$ and may be interpreted as a hyperrational-valued probability function. It assigns an infinitesimal probability to any finite set and a larger probability to any infinite set.

Since algebraic operations on non-standard numbers are equivalent to the termwise application of the corresponding operation on the underlying sequence, from our construction we obtain:

$$P_{num}(A) = \frac{num(A)}{\alpha} = \left[\left\langle \frac{\#(A \cap \{1, \dots, n\})}{n} \right\rangle \right]_{\mathcal{U}} \quad (21)$$

which makes the analogy with asymptotic density complete.

P_{num} is an external object, because its domain is $\mathcal{P}(\mathbb{N})$, which is an external set. This means that the function cannot be obtained directly by taking a standard function and applying the Transfer Principle to it.

5.2 Additivity of the probability function

In this paragraph, we will investigate to what extent our proposal for P_{num} satisfies SUM. We will show that the numerosity function is hypercountably additive (HCA) rather than countably additive (CA). It then follows directly that the probability measure based on it is HCA too.

Addition on ${}^*\mathbb{N}$ First we need to define finite addition on ${}^*\mathbb{N}$. The sum of two hypernatural numbers is defined as the star-map of the standard sum operation (on standard numbers) $+$: ${}^*+$. For any two sequences $\langle a_n \rangle, \langle b_n \rangle$ we have: $[\langle a_n \rangle]_{\mathcal{U}} {}^*+ [\langle b_n \rangle]_{\mathcal{U}} = [\langle a_n \rangle + \langle b_n \rangle]_{\mathcal{U}}$. Since it will be clear from the context whether we are summing standard or non-standard numbers, we may drop the * from the sum-symbol. Likewise, all finite sums are defined.

The countably infinite sum $\sum_{n \in \mathbb{N}}$ is not defined for non-standard terms: since $\sum_{n \in \mathbb{N}} = \lim_{n \rightarrow \infty} \sum_{n=1}^n$ and limits are only defined for standard numbers. Another way to see this is by taking into account that \mathbb{N} is an external set, and summations over such sets are not defined.

An infinite addition that is relevant (always defined) for non-standard numbers is the hyperfinite sum: the equivalence class of a sequence of finite sums. We may also consider the summation over all of ${}^*\mathbb{N}$ —a hypercountable sum $\sum_{N \in {}^*\mathbb{N}}$ — since the latter is an internal set.

num is not CA CA relates a property of the domain of a measure to a property of its range. First of all, the domain has to be a σ -algebra, ensuring that the union of any countably infinite family of sets in the domain is also in the domain. Secondly, countable sums have to be defined on the range. CA then links the two by requiring that the measure of the union of a countable family is equal to the countable sum of measures of the each of the members of the family.

The domain of the numerosity function is $\mathcal{P}(\mathbb{N})$, which is indeed a σ -algebra. The range of the function however is the set of hypernatural numbers, for which the countably infinite sum is undefined. Therefore, the numerosity function cannot be countably additive. The same argument applies to *any* function that maps $\mathcal{P}(\mathbb{N})$ to a set of non-standard values.

num of a sequence of sets So far, the numerosity function is only defined for individual subsets of \mathbb{N} (eq. (17)). Now, we define the numerosity of a sequence of disjoint subsets as follows:

$$\text{num}(\langle A_n \rangle) = {}^*\#({}^*(\langle A_n \rangle) \cap \{1, \dots, \alpha\}) \quad (22)$$

The goal of this paragraph is to find an equivalent form of the above definition, that gives us more insight in the additivity of this function.

First, consider a sequence of (possibly empty) subsets of \mathbb{N} :

$$\langle A_1, A_2, \dots, A_n, \dots \rangle = \langle A_n \rangle_{n \in \mathbb{N}}$$

The star-map of this sequence is a hypersequence of internal subsets of ${}^*\mathbb{N}$:

$${}^*(\langle A_n \rangle_{n \in \mathbb{N}}) = \langle {}^*A_N \rangle_{N \in {}^*\mathbb{N}} \quad (23)$$

with ${}^*A_N = A_N$ for $N \in \mathbb{N}$.

Define the intersection with and the hypercounting function of a hypersequence of subsets of ${}^*\mathbb{N}$ componentwise:

$$(\forall S \in {}^*(\mathcal{P}(\mathbb{N})) \) \quad \langle {}^*A_N \rangle \cap S = \langle {}^*A_N \cap S \rangle \quad (24)$$

$${}^*\#(\langle {}^*A_N \rangle) = \langle {}^*\#({}^*A_N) \rangle \quad (25)$$

By applying equations 23, 24 and 25, the definition given in equation (22) is transformed to:

$$\text{num}(\langle A_n \rangle) = \langle {}^*\#({}^*A_N \cap \{1, \dots, \alpha\}) \rangle \quad (26)$$

Thus, the numerosity of an ω -sequence of sets is a hypersequence of hypernatural numbers; we may refer to its N -th element as $(\text{num}(\langle A_n \rangle))_N$.

num and P_{num} are HCA Now we arrive at the main result of this paper: the numerosity function is hypercountably additive. To prove this, we need to show that for any family of disjoint subsets of \mathbb{N} , $\{A_n \mid n \in \mathbb{N}\}$:

$$\text{num}\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{N \in {}^*\mathbb{N}} (\text{num}(\langle A_n \rangle))_N \quad (27)$$

As an essential step in the proof, we will use that the hyperextension of a countable union of a countable family of sets is equal to the hypercountable

union of the hyperextension of the family (by Transfer, *cf.* [Rubio 1994] p. 104):

$$*\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \bigcup_{N \in {}^*\mathbb{N}} {}^*A_N \quad (28)$$

Proof.

$$\begin{aligned} num\left(\bigcup_{n \in \mathbb{N}} A_n\right) &= {}^*\#\left({}^*\left(\bigcup_{n \in \mathbb{N}} A_n\right) \cap \{1, \dots, \alpha\}\right) \\ &\quad [\text{By definition of num in equation 19}] \\ &= {}^*\#\left(\bigcup_{N \in {}^*\mathbb{N}} {}^*A_N \cap \{1, \dots, \alpha\}\right) \\ &\quad [\text{By equation 28}] \\ &= {}^*\#\left(\bigcup_{N \in {}^*\mathbb{N}} ({}^*A_N \cap \{1, \dots, \alpha\})\right) \\ &\quad [\text{Distributivity of intersection over union} + \text{Transfer}] \\ &= \sum_{N \in {}^*\mathbb{N}} {}^*\#({}^*A_N \cap \{1, \dots, \alpha\}) \\ &\quad [\text{CA of counting function} + \text{Transfer}] \\ &= \sum_{N \in {}^*\mathbb{N}} (num(\langle A_n \rangle))_N \\ &\quad [\text{By equation 26} + \text{remark below it}] \end{aligned}$$

■

Since the numerosity function is HCA, so is the probability function obtained from it (provided that we first define P_{num} of a sequence of sets, as a normalized version of eq. (22)). In particular, for the entire sample space \mathbb{N} , the infinitesimal probabilities of the countably infinite family of singletons do add up to unity. The idea behind the proof is visualized in Figure 1 for the specific case of the countably infinite family of singletons of \mathbb{N} .

It is apparent from the example in Figure 1 that the lottery on \mathbb{N} is HCA in a very specific sense:⁹ for each family, there can always be found a hypernatural number, $K \in {}^*\mathbb{N}$ (α in the example, but possibly larger in other cases), such that the hypercountable sum decomposes in a hyperfinite sum and a hypercountable tail with zero-terms only. Thus we may call num and P_{num} hyperfinitely additive (HFA).

⁹This is completely analogous to the argument, given in section 2.1, showing that a finite lottery is only CA in a trivial sense.

		Hypersequence: << *#(*A _N ∩{1,...,M})>>																
		Sequence: <#(A _n ∩{1,...,m})>																
		(A _n)	*A _N	(m)	M	1	2	3	...	n-1	n	n+1	...	α-1	α	α+1	...	
Family of sets of natural numbers	A ₁ =	*A ₁ = {1}	1	1	1	...	1	1	1	...	1	1	1	...	1	1	1	...
	A ₂ =	*A ₂ = {2}	0	1	1	...	1	1	1	...	1	1	1	...	1	1	1	...
	A ₃ =	*A ₃ = {3}	0	0	1	...	1	1	1	...	1	1	1	...	1	1	1	...
		⋮													⋮			
	A _{n-1} =	*A _{n-1} = {n-1}	0	0	0	...	1	1	1	...	1	1	1	...	1	1	1	...
	A _n =	*A _n = {n}	0	0	0	...	0	1	1	...	1	1	1	...	1	1	1	...
	A _{n+1} =	*A _{n+1} = {n+1}	0	0	0	...	0	0	1	...	1	1	1	...	1	1	1	...
	U _{n∈N}	⋮													⋮			
	N	⋮																
Hyperfamily of sets of hypernatural numbers	*A _{α-1} = {α-1}	0	0	0	...	0	0	0	...	1	1	1	...	1	1	1	...	
	*A _α = {α}	0	0	0	...	0	0	0	...	0	1	1	...	0	1	1	...	
	*A _{α+1} = {α+1}	0	0	0	...	0	0	0	...	0	0	1	...	0	0	1	...	
		⋮													⋮			
	U _{N∈*N}	⋮																
		*Σ _{N∈*N}																
		1 2 3 ... n-1 n n+1 ... α-1 α α+1																

Figure 1: Illustration of the proof for the hypercountable additivity of the numerosity function. In this example, we consider the countably infinite family of singletons of \mathbb{N} , whose union is \mathbb{N} . The numerosity function of a single set requires the extension of sequences (SumCBS) into hypercountably long ones ($*\text{SumCBS}$) and their validation at position alpha. To determine the numerosity of a countably infinite union by addition, the countable family of standard sets has to be extended to a hypercountable family of non-standard sets. By looking at the last horizontal line, we see that the numerosity of $\bigcup_{n \in \mathbb{N}} A_n$ (here $\bigcup_{n \in \mathbb{N}} \{n\} = \mathbb{N}$) is determined as $*\#(*(\bigcup_{n \in \mathbb{N}} A_n) \cap \{1, \dots, \alpha\})$, here $*\#(*\mathbb{N} \cap \{1, \dots, \alpha\}) = @$. In this case, the values in the column with number $M = \alpha$ add up to $@$ too. For each case, a similar table can be made, and the proof states that the same value is always obtained from comparing the α th position of the last row with the hypercountable sum of the α th column.

6 Discussion

In this section, first we evaluate some consequences of our use of free ultrafilters to construct generalized probability functions for infinite lotteries. Subsequently we investigate the relation of a lottery on \mathbb{N} to a different type of infinite lottery: the hyperfinite case. We also compare the non-standard description of a lottery on \mathbb{N} to the best available real-valued approach, that of generalized asymptotic density; we see that SUM is lost in the latter as a result of an accumulation of rounding errors.

6.1 Non-constructiveness

Our approach may be criticized on mildly constructivist grounds. The existence proof for free ultrafilters uses Zorn's lemma and thus depends on the acceptance of the Axiom of Choice. Thus a hyperrational-valued probability measure requires free ultrafilters, which depend on the Axiom of Choice (AC).

To this objection we may reply that the more common solution based on FA, asymptotic density (paragraph 3) requires an extension of the limit concept measure that involves a free ultrafilter too. Therefore, this generally accepted solution is non-constructive too, as has been pointed out by Lauwers [Lauwers 2010]. To those who are unwilling to accept the Axiom of Choice, no measure is available that does any justice to the intuitions underlying fair lotteries. To the rest of us who have no objection to the Axiom of Choice, the HCA, hyperrational probability function is no less acceptable than the FA real-valued one. But there is something to be said in favour of the hyperrational probability function: it gives the SUM-intuition its due.

6.2 Non-uniqueness

Elga has remarked that there are often too many non-standard solutions [Elga 2004]: if we can use any infinite hypernatural number to model a problem, why should we prefer one rather than another? One answer is given by alpha-theory, which develops NSA from the idea of adding a new ideal number, α , to \mathbb{N} . This α can be interpreted as the numerosity of the set \mathbb{N} [Benci & Di Nasso 2003a, p. 357]. Since numerosity theory introduces α as the size of the natural numbers, and constructs the other hypernatural numbers around it, α has a clear interpretation that is hardwired into the theory. It gives us a point of reference among the infinite number of infinite hypernaturals.

We have started from the ultrafilter-construction of ${}^*\mathbb{N}$ and then considered the equivalence class of the sequence $[(1, 2, 3, \dots)]_{\mathcal{U}}$, which we interpret as α (consistent with the axiomatic approach). So, in the present context the accusation of arbitrariness boils down to the choice of a free ultrafilter

\mathcal{U} . A different choice of free ultrafilter produces a different value of α and hence a probability function with the same standard part but infinitesimal differences.

To come back to the *Even* versus *Odd* example: the odd tickets always have a head-start compared to the even ones, for the simple reason that 1 comes before 2. Within our framework, it should not come as a surprise that the weight of this very first ticket may result in an infinitesimal advantage for the whole set of odd tickets. After all, our approach is based on the idea that even in an infinite lottery each ticket has a non-zero probability. For all finite cases with an odd number of tickets, $P_n(\text{Odd}) = \frac{1}{2} + \frac{1}{2n} > \frac{1}{2}$ & $P_n(\text{Even}) = \frac{1}{2} - \frac{1}{2n} < \frac{1}{2}$, whereas $P_n(\text{Odd})$ and $P_n(\text{Even})$ are exactly $\frac{1}{2}$ for all finite lotteries with n even. Since the probability assignment in both cases is different for all finite lotteries, it leads to two different infinite lotteries as well: one in which the odd numbers are in the ultrafilter, such that $P_{num}(\text{Odd}) = \frac{1}{2} + \frac{1}{2\alpha}$ & $P_{num}(\text{Even}) = \frac{1}{2} - \frac{1}{2\alpha}$, and one in which the even numbers are in it, such that $P_{num}(\text{Odd}) = \frac{1}{2}$ & $P_{num}(\text{Even}) = \frac{1}{2}$.

It may look as if we can favour the second solution: a good probability measure should at least respect the limiting frequencies, and they are exactly $\frac{1}{2}$ for *Even* and *Odd*. Take into account however, that limiting frequencies are real-valued, and that our two non-standard solutions have exactly the same standard part, so they are both in accordance with this $\frac{1}{2} / \frac{1}{2}$ solution: limiting frequencies do not rule out one of both.

By observing that $\text{Even} = \{n \in \mathbb{N} \mid n \bmod 2 = 0\}$ and $\text{Odd} = \{n \in \mathbb{N} \mid n \bmod 2 = 1\}$, the above analysis generalizes: for each $m \in \mathbb{N}$, m different scenario's emerge (corresponding to the issue whether $\{n \in \mathbb{N} \mid n \bmod m = k\}$ is in the ultrafilter for either $0, 1, \dots$, or $m - 1$).

It is possible to impose additional constraints on the ultrafilter, such that the set corresponding to $m \bmod 0$ is in the ultrafilter for all $m \in \mathbb{N}$, or equivalently: such that α is a multiple of any finite number [Benci & Nasso 2003b, Mancosu 2009]. However, we see at present no convincing reasons for endorsing any particular constraints of this kind. We already remarked that the infinite lottery violates LABEL. Here we encounter another difference with finite lotteries: the solution is not unique. The problem stated as “Find a probability measure on all of \mathbb{N} that satisfies FAIR, ALL and SUM” is highly underdetermined: there are as many different ways to draw a random number from \mathbb{N} in a fair way as there are free ultrafilters, and the probability function P_{num} given in equation (20) should be seen as a whole family of solutions, all of which are, as far as we can presently see, equally relevant.

6.3 Lotteries on \mathbb{N} versus hyperfinite lotteries

In non-standard measure theory, the emphasis lies on internal measures: measures that can be obtained as the star-map of a standard function. Since

\mathbb{N} is an external set within ${}^*\mathbb{N}$, no internal measure is appropriate to describe the probabilities concerning a lottery over \mathbb{N} .¹⁰ The probability measure we propose in this paper is indeed an external function, but it is closely related to an internal one: the probability function of a hyperfinite lottery.

Consider the finite set $A_n = \{1, \dots, n\}$. Then $[(A_n)]_{\mathcal{U}} = \{1, \dots, \alpha\}$ is a hyperfinite set: an internal set and an initial sequence of ${}^*\mathbb{N}$. Let us consider this hyperfinite set as the sample space of a lottery. The probability measure for this hyperfinite lottery can be obtained by Transfer of the standard measure of a finite lottery, given in equation (1). Thus we obtain the following internal probability measure:

$$\begin{aligned} {}^*P_\alpha : {}^*\mathcal{P}(\{1, \dots, \alpha\}) &\rightarrow {}^*[0, 1] {}^*\mathbb{Q} \\ {}^*A &\mapsto \frac{{}^*\#({}^*A)}{\alpha} \end{aligned} \quad (29)$$

Since this probability measure is obtained as the star-map of a standard probability measure on a finite sample space, which is FA, ${}^*P_\alpha$ is HFA by Transfer.

As such, ${}^*P_\alpha$ cannot be regarded as a satisfactory solution to the problem of denumerably infinite lotteries. The order type of the event space will be vastly different from the order type of \mathbb{N} . Thus, to use terminology introduced earlier, this probability function operates at best with a relabelling of the natural numbers. We have seen in our discussion of LABEL that probability functions on infinite event spaces simply cannot be taken to be invariant under relabelling. So the “internalization” of the problem of infinite lotteries in the non-standard universe does not solve the original problem. Instead, it is a solution to a different problem.

Note that although \mathbb{N} is a subset of $\{1, \dots, \alpha\}$, it is an external one (cf. paragraph 4.6), so it is not part of the domain of ${}^*P_\alpha$. Therefore we cannot use conditionalization on ${}^*P_\alpha$ to define a probability measure on the sample space \mathbb{N} .

If not by conditionalization, what is the relation between the hyperfinite lottery on $\{1, \dots, \alpha\}$ and the hyperrational probability measure on \mathbb{N} ? Any subset A of \mathbb{N} has a hyperextension ${}^*A \in {}^*\mathcal{P}({}^*\mathbb{N})$; the intersection of *A with $\{1, \dots, \alpha\}$ is an internal subset of $\{1, \dots, \alpha\}$ and thus can be assigned a probability value by the internal measure ${}^*P_\alpha$. This is precisely the inner working of P_{num} : a three-step-process that can be read off from equation (19).¹¹

¹⁰With Loeb measures it is possible to transform the non-standard co-domain to a standard one [Cutland 1983], but this still leaves us with a non-standard domain instead of $\mathcal{P}(\mathbb{N})$.

¹¹Our construction of P_{num} can be seen as a specific instantiation of the general suggestion made in [Benci et al 2008, section 5].

6.4 Non-standard probability and asymptotic density

All frequencies that are experimentally accessible concern finite sequences of trials only, and they can be expressed as rational numbers. Limiting frequencies are an infinite idealization of observable frequencies resulting in real-valued probabilities. Our approach is a different idealization that leads to hyperrational probabilities. We will show that its standard part is exactly equal to the limiting frequency. Therefore it is impossible to favour one solution over the other on mathematical or experimental grounds. So we are addressing a question of epistemology: the approaches encapsulate a different vision on probability and it is up to us to decide which one is most in concord with our intuitions.

Co-domain \mathbb{R} versus ${}^*\mathbb{Q}$ For a finite lottery, the probability measure takes values on \mathbb{Q} . However, this co-domain is not closed: a limit of a rational sequence is not always defined in \mathbb{Q} , but it is in \mathbb{R} . In general, the sample space may be infinite; to this end the range is extended to \mathbb{R} . In contemporary presentations, probability theory is usually introduced as a special case of measure theory (e.g. [Dudley 2004]); in this context, the idea that a probability measure is real-valued—which originally was part of an axiom (axiom III in [Kolmogorov 1933, p. 2])—is so basic that it remains a tacit assumption.

Using NSA, the natural extension of the co-domain is ${}^*\mathbb{Q}$ rather than \mathbb{R} . In order to assign sizes to all finite subsets of \mathbb{N} , we already need all of \mathbb{N} (and zero), and to assign probabilities to all finite lotteries, we need all numbers in $[0, 1]_{\mathbb{Q}}$. Thus, in order to assign sizes to all subsets of \mathbb{N} , finite as well as infinite, a larger set is required, which can be obtained using NSA: we consider ${}^*\mathbb{N}$ (and zero), or at least an initial part of it, $\{0, 1, \dots, \alpha\}$ with $\alpha \in {}^*\mathbb{N} \setminus \mathbb{N}$. To assign probabilities to a countably infinite lottery (or all hyperfinite ones), we need $[0, 1]_{*}\mathbb{Q}$ as the range.

Although NSA can be introduced axiomatically, for instance using alpha-theory, and thus need not be intrinsically more difficult than learning standard analysis, the latter was developed first and is still much more common. There are at least two reasons why mathematicians generally prefer the standard framework of classical analysis over the framework of nonstandard analysis. Firstly, nonstandard analysis was rigorously developed much later than standard analysis. So, even though it is not intrinsically more difficult to learn nonstandard analysis, it has the disadvantage of unfamiliarity. Secondly, because of the Transfer Principle, non-standard analysis does not yield any new information about the standard real numbers. However, the concept of a fair infinite lottery simply begs for the co-domain of the sought-for probability functions to be modelled using standard numbers. Thus the Transfer Principle cannot be used here to transport us back to the familiar shore of the standard real numbers. But we do stress that the standard but

ultimately unsatisfactory \mathbb{R} -solution can be interpreted as the standard part of any suitable ${}^*\mathbb{Q}$ -function.

Standard valued approximations and the failure of SUM As we know from section 3, asymptotic density is only FA. Admittedly, CA is not achieved by the ${}^*\mathbb{Q}$ approach either, but the latter tells us why this is not so: the extension of FA for finite lotteries to the countably infinite case does not lead to CA, but rather to HFA (or HCA in a trivial sense). Or, put differently, this solutions shows us that we can extend our finite SUM intuition to HFA. We now investigate how this relates to the FA of asymptotic density.

Let us approximate the hyperreal-valued probability measure (up to infinitesimals) by a real-valued measure. To this end, NSA provides the standard part function, st (paragraph 4.4). First, we will illustrate this by three examples concerning the lottery on \mathbb{N} :

- Single ticket: $A = \{n\}$ for some $n \in \mathbb{N}$. Then $P_{num}(A) = num(\{n\})/\alpha = 1/\alpha$. Since α is an infinite hypernatural, its inverse is an infinitesimal with standard part zero: $st(P_{num}(A)) = 0$.
- All but one ticket. $B = \mathbb{N} \setminus \{n\}$. Then $P_{num}(B) = num(B)/\alpha = (\alpha - 1)/\alpha = 1 - 1/\alpha$. So $st(P_{num}(B)) = 1$.
- Consider an arbitrary $m \in \mathbb{N}$. For all $k \in \{0, 1, \dots, m-1\}$, let C_k be the set $\{n \mid n \bmod m = k\}$. Then $st(P_{num}(C_k)) = st(1/m) = 1/m$ for all k .

In the examples, the standard part of the hyperrational-valued probability measure P_{num} equals the real-valued asymptotic density. This close connection should not come as a surprise, since P_{num} and P_{ad} are, respectively, the ultrafilter equivalence class (or alpha-limit) and the Hahn-Banach limit of one and the same ω -sequence. In fact, what we have observed in the examples holds in general:¹²

$$st \circ P_{num} = P_{ad} \tag{30}$$

Proof.

Fix an arbitrary $A \in \mathcal{P}(\mathbb{N})$. We need to show that $st(P_{num}(A)) = P_{ad}(A)$.

We know that $P_{ad}(A)$ is the Hahn-Banach limit of $\langle a_n \rangle = \langle \frac{\#(A \cap \{1, \dots, n\})}{n} \rangle$ (eq. (3)). From equation (9), we know that in order to proof that $st(P_{num}(A))$ is the Hahn-Banach limit of this sequence, we need to find $H_0, H_1 \in {}^*\mathbb{N} \setminus \mathbb{N}$ with $H_0 < H_1$ and an internal sequence of hyperreals h_{H_0}, \dots, h_{H_1} such that $\sum_{N=H_0}^{H_1} h_N = 1$ and:

¹²This is true for *any* P_{num} in the family of solutions corresponding to the freedom of choice in the free ultrafilter.

$$st(P_{num}(A)) = st\left(\sum_{N=H_0}^{H_1} h_N {}^*a_N\right)$$

First transform the left hand side using equations 6 and 21:

$$P_{num}(A) = \left[\left\langle \frac{\#(A \cap \{1, \dots, n\})}{n} \right\rangle \right]_{\mathcal{U}} = {}^*a_\alpha$$

Then we see that the above equality is equivalent to:

$$st({}^*a_\alpha) = st\left(\sum_{N=H_0}^{H_1} h_N {}^*a_N\right)$$

which holds if we choose $H_0 < \alpha$, $H_1 \geq \alpha$, $h_N = 0$ for all $N \neq \alpha$ and $h_\alpha = 1$.¹³

■

This relation elucidates why (generalized) asymptotic density has poorer additivity properties than the non-standard measure: each time the standard part function is applied, a rounding-off error (up to infinitesimals) is introduced. When adding a finite number of rounded values this still only introduces an infinitesimal error, but when adding an infinite number of approximated terms, the total error may be appreciable. In particular, when summing over the family of all singletons, we see SUM fail most dramatically: the total error is as high as $\alpha \cdot 1/\alpha = 1$, which is 100% error in case of probabilities.

7 Conclusion

In this article we have argued that fair infinite lotteries can best be described using tools and concepts of non-standard analysis. We have constructed a uniform probability measure which is defined on the full power set algebra of \mathbb{N} , and which takes its values in the non-standard extension of the $[0, 1]_{\mathbb{Q}}$ -interval. The construction is closely related to that of asymptotic density measures. The resulting probability measure is uniform, it gives a non-zero probability to the event that a given ticket wins, and it is not just finitely but also infinitely additive.

Lewis argued that every possibility should be assigned a non-zero probability [Lewis 1986]. This has been taken by him and others as a reason for advocating probability functions that take their values in a non-standard $[0, 1]$ -interval. We do not advocate the principle of Regularity in general. But we have seen that in the context of the lottery on the natural numbers, it follows from desiderata that we are committed to on independent grounds.

¹³For the special case where $\langle a_n \rangle$ converges, $st({}^*a_\alpha) = \lim_{n \rightarrow \infty} a_n$ (eqs. (6) and (8)). Thus the proof for that case boils down to: $st(P_{num}(A)) = \lim_{n \rightarrow \infty} a_n = P_{ad}(A)$

So it was natural for us to try to work out a non-standard approach in some detail.

Proposals to construct non-standard measures on infinite sample spaces have been made in the literature [Cutland 1983]. But instead of taking the natural numbers as their domain, these probability functions operate on a non-standard extension of the natural numbers. This entails that they do not provide a genuine solution to the problem of lotteries on the natural numbers; instead, they change the problem. The probability measure that we have constructed is of mixed origins. It stands with one leg in the classical universe, and with its other leg in the non-standard universe.

Asymptotic density and our probability measure are based on the same sequence of partial fractions, and both look at its behavior ‘at infinity’. Only the formalization of this statement is achieved differently: by (a generalization of) classical limits [Schurz & Leitgeb 2008] in the first case, and by free ultrafilter-based equivalence classes in the second. Asymptotic density takes values in the standard $[0, 1]$ -interval. The price for this is that such probability functions can only be finitely additive for fair lotteries on \mathbb{N} .

The solution that we propose meets the conceptual and intuitive requirements connected with a lottery on \mathbb{N} : it is fair, uniform, defined on the whole power set algebra of \mathbb{N} , and infinitely additive. We have seen that we do not end up with a unique solution. The probability functions that we propose are only determined up to the choice of an ultrafilter. As far as we can presently see, no choice of ultrafilter is superior to any other. We do not exclude that when more intuitive constraints are imposed, the class of satisfactory probability functions can be narrowed down further. But at present we see no way of doing so.

The most notable feature of our solution is that whereas it is infinitely additive, its additivity behavior is not adequately described by summing over the natural numbers, but by summing over a non-standard extension of the natural numbers. One might naively expect the relevant sum generally to be an ω -sum of non-standard weights. But it emerged that the probability of an ω -length family of events will be a sum of a much longer order type. In this context, it should be recalled that even Kolmogorov, the discoverer of the classical σ -additivity axiom, emphasized that the condition is merely a useful assumption of idealization but is not contained in the meaning of probability [Kolmogorov 1933, p. 15].¹⁴ It turned out that in the context of lotteries on the natural numbers, σ -additivity is not the *right* idealization.

The obvious next step would be to consider larger, nondenumerable lotteries, such as lotteries on \mathbb{R} . The description of fair lotteries on \mathbb{R} is a

¹⁴Here, Kolmogorov writes: “Infinite fields of probability occur only as idealized models of real random processes. *We limit ourselves, arbitrarily, to only those models which satisfy Axiom VI.*” (Emphasis in the original.) Here, ‘field’ means ‘algebra’ and ‘Axiom VI’ refers to the ‘Axiom of Continuity’ which, together with FA, automatically leads to CA in cases where the algebra is a σ -algebra.

mathematically significantly more difficult affair, and will be left for future work.

References

- [Bartha 2004] Bartha, P. *Countable additivity and the de Finetti lottery*. British Journal for the Philosophy of Science **55** (2004) p. 301–321.
- [Benci & Di Nasso 2003a] Benci, V., Di Nasso, M. *Alfa-theory: an elementary axiomatic for nonstandard analysis*. Expositiones Mathematicae **21** (2003) p. 355–386.
- [Benci & Nasso 2003b] Benci, V., Di Nasso, M. *Numerosities of labelled sets: a new way of counting*. Advances in Mathematics **173** (2003) p. 50–67.
- [Benci et al 2006a] Benci, V., Di Nasso, M., Forti, M. *The eightfold path to nonstandard analysis*. In: Cutland, N. J., Di Nasso M., Ross, D. A. (eds.) *Nonstandard methods and applications in mathematics. Lecture notes in logic, Vol. 25*. Association for Symbolic Logic, AK Peters, 2006, p. 3–44.
- [Benci et al 2006b] Benci, V., Di Nasso, M., Forti, M. *An Aristotelian notion of size*. Annals of Pure and Applied Logic **143** (2006) p. 43–53.
- [Benci et al 2008] Benci, V., Galatolo, S., Ghimenti, M. *An elementary approach to stochastic differential equations using infinitesimals*. unpublished manuscript.
- [Cutland 1983] Cutland, N. *Nonstandard measure theory and its applications*. Bulletin of the London Mathematical Society **15** (1983) p. 529–589.
- [de Finetti 1974] de Finetti, B. *Theory of probability*. Vols. 1 & 2, translated by A. Mach and A. Smith. Wiley, 1974.
- [Di Nasso & Forti 2010] Di Nasso, M., Forti, M. *Numerosities of point sets over the real line*. Transactions of the American Mathematical Society, in press (2010). DOI: 10.1090/S0002-9947-2010-04919-0.
- [Douven et al 2010] Douven, I., Horsten, L., Romeijn, J.-W. *Probabilist anti-realism*. Pacific Philosophical Quarterly **91** (2010) p. 38–63.
- [Dudley 2004] Dudley, R. M. *Real analysis and probability*. Cambridge University Press, 2004.
- [Easwaran] Easwaran, K. *Regularity and infinitesimal credences*. Article in preparation.

- [Elga 2004] Elga, A. *Infinitesimal chances and the laws of nature*. Australasian Journal of Philosophy **82** (2004) p. 67-76
- [Hájek] Hájek, A. *Staying regular*. Article in preparation.
- [Kolmogorov 1933] Kolmogorov, A. N. *Grundbegriffe der Wahrscheinlichkeitrechnung*. Ergebnisse Der Mathematik, 1933. Translated by Morrison, N. *Foundations of probability*. Chelsea Publishing Company, 1956 (2nd ed.).
- [Komjáth & Totik 2008] Komjáth, P., Totik, V. *Ultrafilters*. American Mathematical Monthly **115** (2008) p. 33-44.
- [Kubilius 1962] Kubilius, J. *Probabilistic methods in the theory of numbers*. 1962. Translated by Burgie, G. and Schuur, S., *Translation of mathematical monographs*. Vol. 11. American Mathematical Society, 1978 (3rd ed.).
- [Lauwers 2010] Lauwers, L. *Purely finitely additive measures are non-constructible objects*. Working paper, unpublished (2010). Available from URL: <http://www.econ.kuleuven.be/ces/discussionpapers/Dps10/DPS1010.pdf>, Retrieved: May 29 2010.
- [Lavine 1995] Lavine, S. *Finite mathematics*. Synthese **103** (1995) p. 389-420.
- [Lewis 1980] Lewis, D. *A subjectivist's guide to objective chance*. In: Jeffrey, R. (ed.) *Studies in inductive logic and probability*. Vol. II. University of California Press, 1980.
- [Lewis 1986] Lewis, D. *Philosophical Papers: Volume II*. Oxford University Press, 1986.
- [Mancosu 2009] Mancosu, P. *Measuring the size of infinite collections of natural numbers*. The Review of Symbolic Logic **2** (2009) p. 612-646.
- [McCall & Armstrong 1989] McCall, S., Armstrong, D. M. *God's lottery*. Analysis **49** (1989) p. 223-224.
- [Ramsey 1926] Ramsey, F. P. *Truth and probability*. 1926. In: Braithwaite, R. B. (ed.) *Foundations of mathematics and other essays*. Routledge & P. Kegan, 1931, p. 156-198.
- [Robinson 1966] Robinson, A. *Non-standard analysis*. North-Holland Publishing Company, 1966.
- [Rubio 1994] Rubio, J. E. *Optimization and nonstandard analysis*. Marcel Dekker, 1994.

- [Schurz & Leitgeb 2008] Schurz, G., Leitgeb, H. *Finitistic and frequentistic approximation of probability measures with or without σ -additivity*. *Studia Logica* **89** (2008) p. 257–283.
- [Skyrms 1980] Skyrms, B. *Causal Necessity*. Yale University Press, 1980.
- [Tenenbaum 1995] Tenenbaum, G. *Introduction to analytic and probabilistic number theory*. Cambridge Univ. Press, 1995.
- [Truss 1997] Truss, J. *Foundations of mathematical analysis*. Clarendon Press, 1997.
- [Väth 2007] Väth, M. *Nonstandard analysis*. Springer, 2007.
- [Williamson 2007] Williamson, T. *How probable is an infinite sequence of heads?* *Analysis* **67** (2007) p. 173–180.